

# 多视角篮球数据集

---

## Step0-数据集介绍

---

### 数据集概述

此数据集旨在用于预测篮球比赛中哪个球员最终会进行投篮动作。数据集包含了不同球员在比赛中的投篮视频片段，并经过裁剪、切割和标注处理，以便作为机器学习或深度学习模型的训练数据。

### 标注任务概述

1. 数据集规模：拟构建中等规模数据集。收集约5000段视频片段，总时长大约7个小时，每段视频3-10秒不等。对每段视频进行关键帧的选取及标注，最后得到的所有数据集文件大小为120G左右。
2. 标注流程
  1. 从原始视频素材中，筛选得到画面中包含全体球员的可用视频片段，对视频片段进行剪切保存。
  2. 根据视频片段中的信息，对得到的视频片段进行规范化的重命名。
  3. 对视频片段进行切帧，选取需要的关键帧进行标注。

### 视频素材概述

视频素材为每一天的NBA十佳球/五佳球集锦，每一球均为一个片段。我们需从中筛选出可用的视频片段。同时，每一天通常还包括最佳镜头、最佳扣篮、最佳助攻各一个，也分别为一个片段，也需要进行筛选。需要注意，“最佳”片段中可能含有之前的十佳球/五佳球中已经出现过的片段，则此时应注意去重。

### 标注工具

1. 视频片段提取工具：LosslessCut（见附件1）
2. 视频片段切帧工具（见附件2）
3. 图片标注工具：Darklabel（见附件3）